DOCUMENT RESUME

ED 042 809 TM 000 092

AUTHOR Hall, Marian D.; And Others
TITLE An Experimental Assessment Program in an Inner City School.
PUB DATE Mar 70
NOTE 16p.; Paper presented at the annual meeting of the American Educational Research Association, Minneapolis, Minneaota, March, 1970

EDRS PRICE EDRS Price MF-$0.25 HC-$0.90
DESCRIPTORS Auditory Visual Tests, *Classroom Environment, Classroom Observation Techniques, Disadvantaged Youth, Educational Testing, Grade 1, *Inner City, Interaction, *Learning Disabilities, Learning Motivation, Learning Processes, Predictor Variables, *Student Ability, Student Behavior, *Student Teacher Relationship, Teacher Behavior, Urban Education

ABSTRACT

This study is an effort to assess both the internal and external conditions of learning. The sample included 99 first graders in four classrooms of one inner city school in a large midwestern city. The assessment measure was a battery of tests selected from frequently used tests of visual perception, auditory discrimination, language, memory, cognition, and motor skills. The inner city classrooms, when compared to standardized populations, showed much greater variability and significant means differences on the majority of test items. The patterns of disability fell into modal profiles that presented essential information for educational prescription. Behavioral observations of dependency and aggression, when combined with ratings of teacher-pupil interaction, yielded information about classroom climate. Short term gains from the study included modification of curriculum grouping and staffing. Long term gains are expected to include progress toward validating an assessment battery and supplemental knowledge about the interaction between the pattern of children's capabilities and the climate of the classroom. (Author/AE)

ED042809

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

An Experimental Assessment Program in an Inner City School

Marian D. Hall
Elizabeth M. Hyde
Georganne R. Shove

University of Minnesota

Paper read at the Annual Meeting of the American Educational Research Association

1970

TM 000 092

Psychologists have long concentrated their efforts on assessing the performance of children, analyzing and labeling intellectual abilities and personality configurations. With equal assiduousness, educators have concentrated on task analysis, parsing the content of the school curriculum. Neither approach, when used alone, has prevented the escalation of numbers of children referrred for learning disabilities nor presented school personnel with maximally effective predictive and remedial instruments. Studies that attempt to evaluate the best method of teaching reading without matching child to method leave the front-line practitioner, the teacher, with many unanswered questions. Likewise, the traditional psychological report that presents a detailed analysis of an individual child's motivational state and capabilities but fails to consider relevant task variables has proved to be a hollow substitute for effective prescription.

The study reported here is a preliminary attempt to move toward convergence of the task analysis and the child evaluation models. It takes its point of view from Gagne' (1965), emphasizing both the internal and the external conditions of learning, i.e., the previously learned capabilities and motivations on the one hand and the stimulus situation, the directions and feedback, on the other. The two chief foci of attention of the investigation were the exploration of ways of charting profiles of pupil capabilities in those areas presumed to subserve learning to read and the identification of certain situational variables that affect the external conditions of learning, namely the interaction between child behaviors (dependency and aggression) and teacher behaviors (direct and indirect influence categories).

Not only do individual children show a wide range of diverse abilities, but the climate of the classroom perforce multiplies these diverse patterns. This is especially true in inner-city schools and neighborhoods in transition. The school situation that was the scene of this study had rapidly changed character in the previous five to ten years, shifting from a predominantly middle-class area to a working class neighborhood with a high proportion of minority group families, many newly arrived in the state. The school, accordingly, was raising essential questions about the validity of the criteria, goals, and expectancies that had served them before, and rightfully treating as suspect the test measures and teaching procedures appropriate to the needs of a previous community.

The Study: Assessing Pupil Capabilities and Classroom Climate

The Sample

The project involved 99 children in the four first-grade classrooms of one inner-city school in a large Midwestern city. The student enrollment in this school is 70% Negro, being drawn from both "project" housing areas, single rental units, and family-owned units.

The Measures

The immediate goal of the project was to construct a test battery that would, within feasible limitations of examiner's time and skill, allow grouping of children according to modal profiles of learning disability. The relevant dimensions were assumed, on the basis of both theoretical and empirical reasoning, to include: visual perception, auditory discrimination, language, memory, cognition, motor skills (Shove, 1968).

In addition, behavior samplings were made on a randomly selected group of five boys and five girls from each of the four classrooms to determine the degree and type of dependency and aggression behaviors and the classroom antecedents of such behaviors (Hyde, 1968). The four teachers were also observed without their knowledge during four five-minute periods and their behaviors classified into the ten categories of teacher-child interaction proposed by E. J. Amidon and N. A. Flanders (1963).

Areas of assessment and tests used were as follows:

I Intellectual-Achievement
  Peabody Picture Vocabulary
  Kindergarten Metropolitan Achievement Tests

II Emotional-Social-Cultural
  Lambert and Bower, "A Process of In-School Screening of Children with Emotional Handicaps"
  Behavior Sampling
  Parent Questionnaires
  WISC Comprehension (1-5)

III Attentional Factors
  Behavior Sampling

IV Perception and Discrimination
  Visual - Frostig Developmental Test of Visual Perception
  Auditory - Wepman Auditory Discrimination Test

V Assimilation
  Integration of Modalities - Frostig; Lincoln-Oseretsky Motor Development Scale; Stanford-Binet VI-6 (maze-tracing); ITPA auditory-vocal sequencing; ITPA auditory-vocal automatic
  Organization of received information - Stanford-Binet VI-4 (number concepts); Stanford-Binet VI-2 (differences)

VI Memory
  Stanford-Binet (M) VII-2 (memory for sentences II); ITPA auditory-vocal sequencing

VII Analysis
  ITPA auditory vocal association

VIII Expression
  Motor - Lincoln Oseretsky; Binet VI-6 (maze tracing)
  Vocal - ITPA auditory-vocal automatic; ITPA auditory-vocal association

Results: Assessing Pupil Capabilities

The test scores were analyzed by computer to obtain means and standard deviations of each test administered as well as correlations between the various tests. This analysis consisted of twelve separate measures. The Stanford-Binet and WISC items were not included because the pass-fail scoring provided no distribution on which to compute deviations. Only those subjects for whom data were complete were included (N=81). T-tests were computed to test the level of significance of the difference between mean scores of the sample and the standardization population of the test. This was not possible for the Wepman or the Lincolm Oseretsky due to lack of normative data in the test manuals.

Insert Table 1 about here

Thus it will be seen that the mean differences are highly significant, greater than $p. < .0005$ in all but one case, indicating major differences between the two populations.

The ranges of scores on each test are shown on Table 2.

Insert Table 2 about here

The four items from the Stanford-Binet and the WISC Comprehension items were combined to measure a "cognitive ability factor, in terms of Bateman's intermediate process". Three of the Binet subtests (Maze Tracing, Differences, Number Concepts) are from the six-year level, while one (Sentence Memory) is from the seven-year level since no test of word memory existed at the six-year level. A raw score of six on WISC Comprehension is the mean score (scale score = 10) for the norm age group 6-4 through 6-7 (Wechsler, 1949) and was considered the "pass" score in this sample. Out of the total of five different tests in the cognitive area, the mean number of items passed by the subject population was 3.17. The percentage passing each item, compared with per cent passing in the norm group is shown on Table 3. It should be noted that since Sentence Memory II is from the 1937 Form M Binet, per cent of the norm group passing is not precisely mentioned. The figure given is the <u>average</u> per cent passing Form M items in the 1960 standardization (C.A. six) (Terman & Merrill, 1960, p. 32).

Insert Table 3 about here

The various tests were intercorrelated with results and degrees of significance as shown on Table 4.

Insert Table 4 about here

The Results: Assessing Classroom Climate

Three aspects of classroom climate were measured: aggressive and dependent behaviors in children were rated on a direct observational schedule (Hyde, 1968) carefully constructed to reflect overt behaviors of first-grade populations such as the one under study; teacher perception, peer perception, and self-perception were measured according to the instruments in the <u>Process for In-School Screening of Children with Emotional Handicaps</u> (Lambert & Bower, 1961); and teacher's classroom behavior was observed and classified into the ten categories listed in the Flanders & Amidon (1963) matrix system. Each of the four teachers was rated on three different occasions while introducing typical classroom activities. None of the teachers was aware she was being rated.

Analysis of this data (Table 5) showed differential effects among the four teachers across the nine categories of aggressive and dependency behaviors for the two sexes. Differences among teachers appear to be due to the boys; unlike

---

Insert Table 5 about here

---

many previous studies, however, the frequency of higher scores for boys holds for dependent as well as aggressive behaviors. Further contrasts revealed that peer aggression was significantly higher (p .01) than teacher aggression and teacher dependency significantly higher ($p < .01$) than peer dependency. The patterning of these behaviors according to classroom is seen in Figure 1.

Insert Figure 1 about here

Ratings of teacher perception, peer perception, and self-perception show no significant differences across classrooms or between sexes. Product-moment correlations among these three interpersonal ratings and the aggression and dependency measures show that teacher ratings and peer ratings correlate at .404 ($p < .05$) and teacher ratings correlate positively and significantly with aggression ($r = .528$, $p < .01$), but are not related to any form of dependency, suggesting that aggressive students were rated negatively by their teachers whereas dependent students were not.

The observations of teacher classroom behavior was subjected to a matrix analysis, resulting in a 4x10 table, consisting of the total percentage of time each of the four teachers spent in the ten categories. A Friedman Two-Way Analysis of Variance resulted in a X^2 value of 20.1545 ($p < .05$), suggesting that significant differences did exist among the four teachers in their distribution of behaviors among the ten categories. Figure 2 illustrates the patterning of such differences.

Insert Figure 2 about here

Summary and Discussion of Findings

1) On all tested skills, the four classrooms in this inner-city school showed much greater variability than is normally expected. With an average chronological age of 6-6, the age equivalents for the six skill areas ranged from 2-3 years at the lower extreme to 8-10 years at the upper. This has two significant aspects. First, it should be noted that the deviation from the expected performance was not solely due to substandard scores, and that in every skill area and in every classroom there were children achieving at or above age expectation. (In fact, as a subsequent item will show 23 of the 99 were at or above age expectation on all skills.) In the second place, the impact of this immense spread on the teaching task cannot be ignored; the demands made upon teachers' skill, energy, and resourcefulness are vastly greater than under usual circumstances.

2) The primary deficit areas for the group were in the related auditory perception and language skills. This finding agrees with the work of other investigators of inner-city elementary education (Whiteman and Deutsch, 1968; Bloom, Davis, and Hess, 1965; Bereiter and Engelmann, 1966; Gray and Klaus, 1965). On the Wepman Auditory Discrimination measure, these children at 6-6 performed at levels less adequate than the normative population of five year olds. An interesting side finding was a reliability check made to test the hypothesis that poor auditory discrimination was related to the difference between speech patterns of professionals (teachers, psychologists, etc.) and the prevailing neighborhood speech. N tester effect was found to be operating generally; however, it is not to be ignored that 19% of the children showed measurable improvement when the words were presented by a neighborhood aide as compared to a professional tester, indicating that either or both differential pronunciation or motivational factors may be affecting performance.

The language measures included the Peabody Picture Vocabulary, a test of vocabulary comprehension, and two subtests of the Illinois Test of Psycho-linguistic Abilities, (The Auditory-Vocal Automatic and the Auditory-Vocal Association subtests, which are presumed to measure grammatical and analogic use of language). On the Peabody the children scored at the 28th percentile for their age group, for a language quotient of 89. The subtests of the ITPA yielded mean group age equivalents of 5-9 and 5-10, and were the tests that most frequently differentiated this population from the normative group, approximately half of the children showing significant deficit in the area of language usage, grammatical forms (verb forms, plurals, etc.) and language concepts (e.g., "I hit with my hands; I kick with my____"). Both of these tests have been found to be directly influenced by cultural background and to be especially resistant to the general-enrichment-of-program approach. Hence, such programs as Head Start have had little impact on this form of language deficit, and it was not possible to differentiate language-efficient from language-deficient children on the basis of pre-school training.

3) The tests of visual perception (Frostig) showed a great deal of scatter, with the more general perceptual tasks (Eye-Motor Coordination and Figure-Ground Discrimination) generally accomplished well above the expected level by inner-city pupils and the performance on tasks that call for visual discrimination within a narrower field and more directly related to reading skills (Form Constancy, Position in Space, and Spatial Relationship) below age expectancies. In the two areas of superior achievement, the average for the group was at levels appropriate for 7-9½ year olds. On these tasks, however, there was a great deal of variability within the population and some children performed as low as the 2 and 3 year levels. In the three tasks that averaged below the norm, the mean scores for the group approximated those for the normative population at 6-0. The significance of this overall deficit is that in ability to recognize a letter form in various contexts or positions in space (reversals, rotations), or relationships (sequencing), the average inner-city child after six months in first-grade was performing much like the average child in the standardization sample had upon entering first grade.

4) The memory tasks indicate that rote memory is, for the group, close to expected levels and that when memory deficit occurs it may well be that it is confounded with a related language deficit. For instance, on the Memory for Sentences task, there was a tendency to lump words together, leave out connectives,

or delete verb endings. It seems likely that such errors belong to a common factor that also includes Auditory Discrimination and Language Usage.

5) The tests of Comprehension and Cognition (subtests such as Maze-tracing, Differences, and Number Concepts from the S-B and Comprehension from the WISC) indicated that, in general, the inner-city children could perform in areas that called for common-sense and problem-solving skills as well as children of their chronological age may be expected to perform and in some instances (Maze-tracing and Differences) the percent of the students passing the items was higher for this population than for the normative group, in this instance the Stanford-Binet standardization sample.

At this juncture, it is necessary to point out that it would be desirable to replicate this kind of screening procedure or in other ways to develop a local normative sample. For the present, comparisons are based on the reported standardization data, all white and middle class samples. Not infrequently, the test constructors have furnished incomplete information on sample means and standard deviations. Furthermore, some of the tests used in the battery (ITPA, Frostig) were still in experimental stages. They are, however, among the best instruments currently available and the need to press toward a total and efficient screening battery determined their use. Further factor analytic studies are needed at this point.

6) This information, plus a rough, five-point ranking of all the children on scales of Motor Coordination (Lincoln-Oseretsky Tests of Motor Skills) and also ratings on Dependency and Aggression (based on behavior sampling) was entered onto individual profiles for each child. Profile analysis and sorting led to the following grouping:

23 children showed no deficiency areas
22 children showed only language deficiency
19 children showed primary language deficiency and secondary visual perceptual deficiency
8 children showed primary visual perceptual and secondary language deficiency
11 children showed visual perceptual deficiency only
2 children showed memory deficiency and visual perceptual deficiency
5 children showed general mental retardation
9 children showed primarily emotional problems

99 total profiles

7) The behavioral observations of dependency and aggression were obtained in a carefully designed and implemented study. Five categories of aggression (verbal-peer, physical-peer, direct-teacher, subtle-teacher, and indirect-peer) and four categories of dependency (verbal-peer, physical-peer, direct-teacher and subtle-teacher) were tallied on a time-sampling basis. Significant sex differences were found, with the first grade boys higher than girls on the aggression scales ($p < .01$) and, unlike many previous studies, also on the dependency scales ($p < .05$). More revealing, however, were differences on these scales according to classroom teacher and correlations between the dependency and aggression behaviors of the children with teachers' ratings of the children on the Bowers and Lambert Attitude Measures, children's self-ratings, and observations of teacher behaviors based on the Amidon-Flanders direct-indirect

influence categories. Direct teacher influence includes such behaviors as lectures, gives directions, criticizes, and justifies authority; indirect teacher influence includes such behaviors as accepts feelings, praises or encourages, accepts or uses ideas of students, asks questions. The one teacher who showed markedly less use of indirect influence (only 1/5 as much as the others) and who generally showed a high degree of authoritarian control, critized a great deal and permitted very little student participation, had markedly less aggressive verbal behavior in her classroom, directed either at teacher or peers. She also had the least amount of verbal peer dependency. The other three teachers, who displayed similar patterns of indirect influence behaviors, had generally more active classrooms, more interactions between pupils and teacher were taking place, both of the aggressive and need-satisfying kinds. The quality of control of the one teacher who used primarily direct influence methods was such that children turned not to her but to eachother for need satisfaction. It is noteworthy that the self-ratings (Bowers and Lambert) for this group indicates the lowest self-concept of all the four classes.

Implications for programming

Specific suggestions that follow from studies such as this are aimed at meeting immediate needs: changes in grouping, team teaching, remedial curriculum, supplementary staffing. More significant for long-range gains, however, are suggestions related to continued and refined investigation. The validation of compact batteries of test items that will allow us to graph profiles of mental abilities across varying developmental stages and complex types of learning is a major priority if we are to make headway in solving what has been aptly termed "the problem of the match". More sophisticated measures of classroom interaction, particularly more specificity about the varieties of child behaviors, will require new measures similar to the time and event sampling reported here for dependency and aggression in the classroom. And, most importantly, longitudinal follow-up data on the interaction between the internal capabilities, the store of prerequisite knowledge,and the external conditions, the instructional climate, is necessary before we can do more than mouth vague generalities about intervention and prescription. Some hint of the complexity of the task before us and, perhaps, some explanation for the current failure to find easy solutions becomes apparent when we keep in mind that we are searching for the interface between assessment and prescription, the confluence of learning and instruction, the optimum match between a child and a method.

References

Amidon, E. J. and Flanders, N. A. The role of the teacher in the classroom: a manual for understanding and improving teachers' classroom behavior, Minneapolis, Minn., Amidon and Associates, Inc., 1963

Bereiter, C. and Engelmann, S. Teaching disadvantaged children in the preschool. Englewood Cliffs, New Jersey: Prentice-Hall, 1966

Bloom, B. S., Davis, A., and Hess, R. Compensatory education for cultural deprivation. New York: Holt, Rinehart & Winston, 1965

Bower, E. H. Early identification of emotionally handicapped children in school. Springfield, Ill: Chas. C. Thomas, 1960

Gagne' R. M. The conditions of learning. New York: Holt, Rinehart & Winston, 1965

Gray, S. W., and Klaus, R. A. An experimental preschool program for culturally deprived children. Child Development, 36, 4, 887-898.

Hyde, E. Aggression and dependency in first grade students and corresponding teacher classroom behavior. (unpublished paper, University of Minnesota,) Minneapolis, Minnesota, 1968

Lambert, N. M. and Bower, S. M. A process for in-school screening of children with emotional handicaps: a manual. Sacramento, Calif.: Calif. State Dept. of Educ., 1961

Shove, G. Toward a diagnosis of learning disability. (Unpublished paper, University of Minnesota) Minneapolis, Minn., 1968

Whiteman, M., Brown, B. R., and Deutsch, M. Social disadvantage as related to intellective and language development. In Deutsch, M., Katz, I., and Jensen A. R. Social class, race, and psychological development, New York: Holt, Rinehart & Winston, 1968

Table 1

Test Score Differences on Learning Disability Battery
between Inner-City School First Graders (N=81)
and Normative Population

Variable No.	Test	Mean	Standard Devia.	Age Equiva.	Norm Mean (Ca6½-7)	t
1	ITPA &-voc automatic	11.2	3.57	5-9	13.5	-5.77**
2	ITPA &-voc association	16.0	3.81	5-10	18.8	-6.59**
3	ITPA &-voc sequencing	22.1	6.52	6-3	24.0	-2.61*
4	Lincoln-Oseretsky	25.2	8.80	---	---	---
5	Frostig I	19.7	5.30	9-6	9.93	+≪16**
6	Frostig II	15.2	4.19	7-0	10.09	+10.89**
7	Frostig III	7.0	3.36	6-0	9.78	-7.40**
8	Frostig IV	6.1	1.62	6-3	10.15	-≪22**
9	Frostig V	4.3	1.75	6-0	10.36	-30.99**
10	Peabody Picture Vocabulary	55.0	6.81	5-11	60.79	-7.62**
11	Wepman I	7.0	4.00	>5-0	---	---
12	Wepman II	7.0	3.91	>5-0	---	---

* = .01 ** = .0005

Table 2

Ranges of Scores of Inner-City First-Graders (N=81)

on Twelve Items in Learning Disability Battery

Variable Number	Test Given	High Score	High Age Equiv.	Low Score	Low Age Equiv.
1	ITPA &-voc automatic	18	8-4	1	2-4
2	ITPA &-voc association	24	>9-0	5	3-1
3	ITPA &-voc sequencing	40	>8-6	6	2-9
4	Lincoln-Oseretsky	41	---	3	---
5	Frostig I	32	10-0	4	3-9
6	Frostig II	20	8-3	0	2-6
7	Frostig III	14	9-0	0	2-6
8	Frostig IV	8	8-9	1	3-3
9	Frostig V	7	8-3	0	4-0
10	Peabody Picture Vocabulary	77	10-4	37	3-8
11	Wepman I	0	---	15	---
12	Wepman II	1	---	15	---
X	Binet & WISC	5	>7-0	0	<5-0

Table 3

Percent of Inner-City First Graders

Passing Selected Cognitive Ability Items

Item	Subject Percent Passing	1960 Norms % Passing
Maze Tracing (VI,6)	76	82
Differences (VI,2)	84	61
Number Concepts (VI,4)	69	71
Sentence Memory (form M, VII,2)	30	41
WISC Comprehension (1-5)	57	--

Table 4

Intercorrelation Matrix of Items in Learning Disability Battery

Variables	1	2	3	4	5	6	7	8	9	10	11	12
					N = 81							Wepman II N = 42
1 ITPA Aud-Voc Auto			*** .59	** .26	* .24	.03	* .20	*** .33	*** .33	*** .28	*** .59	*** -.44	* -.26
2 ITPA Aud-Voc Assoc.				*** .35	*** .33	.06	*** .37	*** .48	*** .51	*** .48	*** .52	*** -.45	** -.36
3 ITPA Aud-Voc Seq.					*** .31	-.01	.09	.17	.12	* .27	.16	*** -.41	-.06
4 Lincoln-Oseretsky						.18	* .25	*** .29	* .21	*** .32	.08	* -.21	-.01
5 Frostig I							** .27	** .27	*** .32	*** .33	.17	.07	*** -.43
6 Frostig II								*** .42	*** .35	*** .37	.00	-.14	-.12
7 Frostig III									*** .43	*** .28	** .26	*** -.38	* -.33
8 Frostig IV										*** .45	*** .44	*** -.32	* -.35
9 Frostig V											* .25	*** -.29	-.06
10 Peabody												* -.19	-.10
11 Wepman I													*** -.48
12													

* = .05
** = .01
*** = .005

Table 5

Summary Table for Aggressive and Dependent Behaviors

of Inner City School First-Graders (N=40)

Source	Degrees of Freedom	Sum of Squares	Mean Square	F Ratio	Level of Significance
Classroom	3	1169.900	389.967	12.739	.01
Sex	1	751.111	751.111	24.536	.01
Classroom X Sex	3	814.934	271.645	8.874	.01
Error	32	979.599	30.613	---	
Behavior	8	5234.589	654.324	14.811	.01
Behavior X Classroom	24	3414.900	142.288	3.221	.01
Behavior S Sex	8	1335.489	166.936	3.779	.01
Behavior X Classroom X Sex	24	1995.866	83.161	1.882	.01
Error	256	11309.601	44.178	---	
Total	359	27005.989			

10 Squares to the Inch

NATIONAL 12-280 Made in U.S.A.

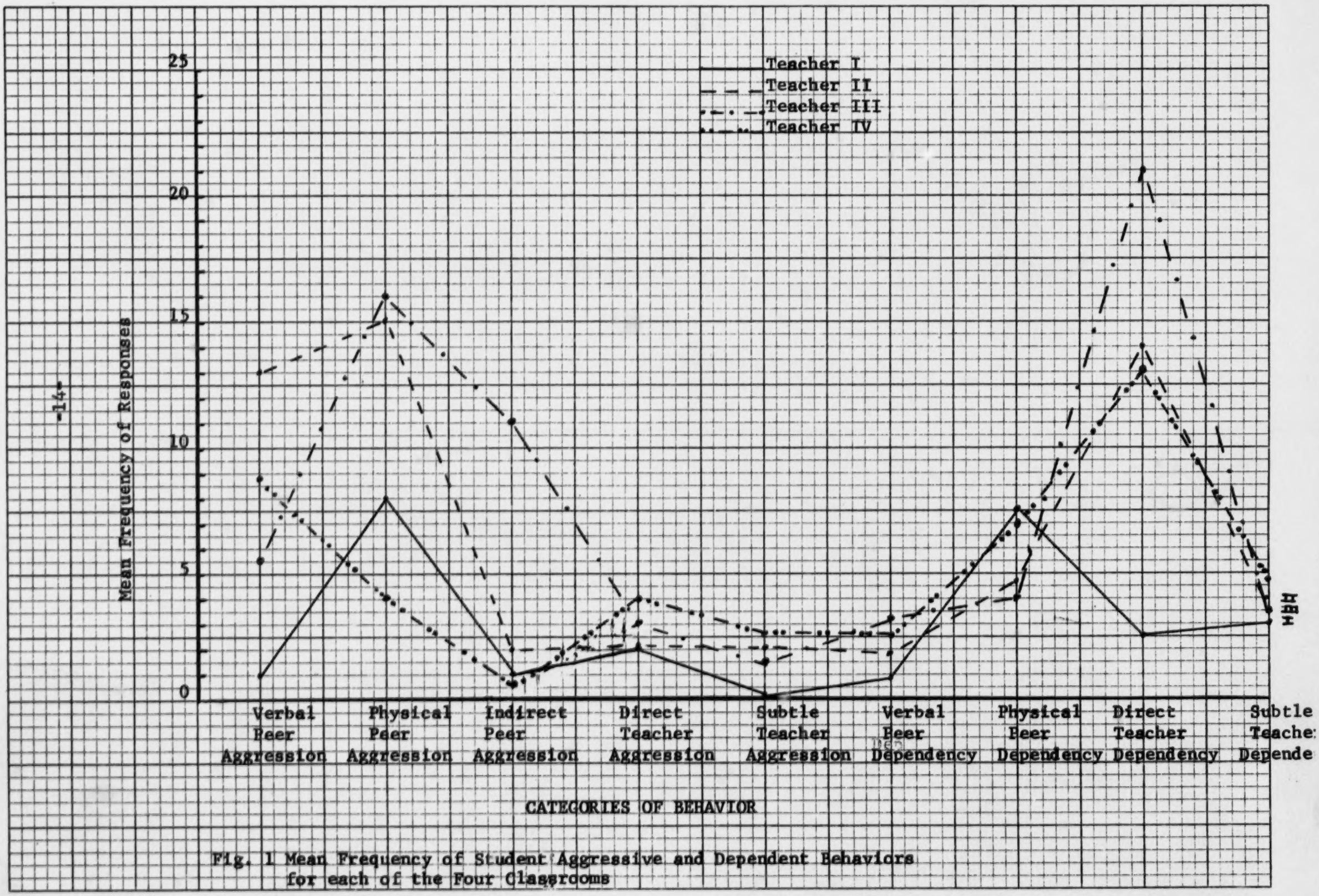

Fig. 1 Mean Frequency of Student Aggressive and Dependent Behaviors for each of the Four Classrooms

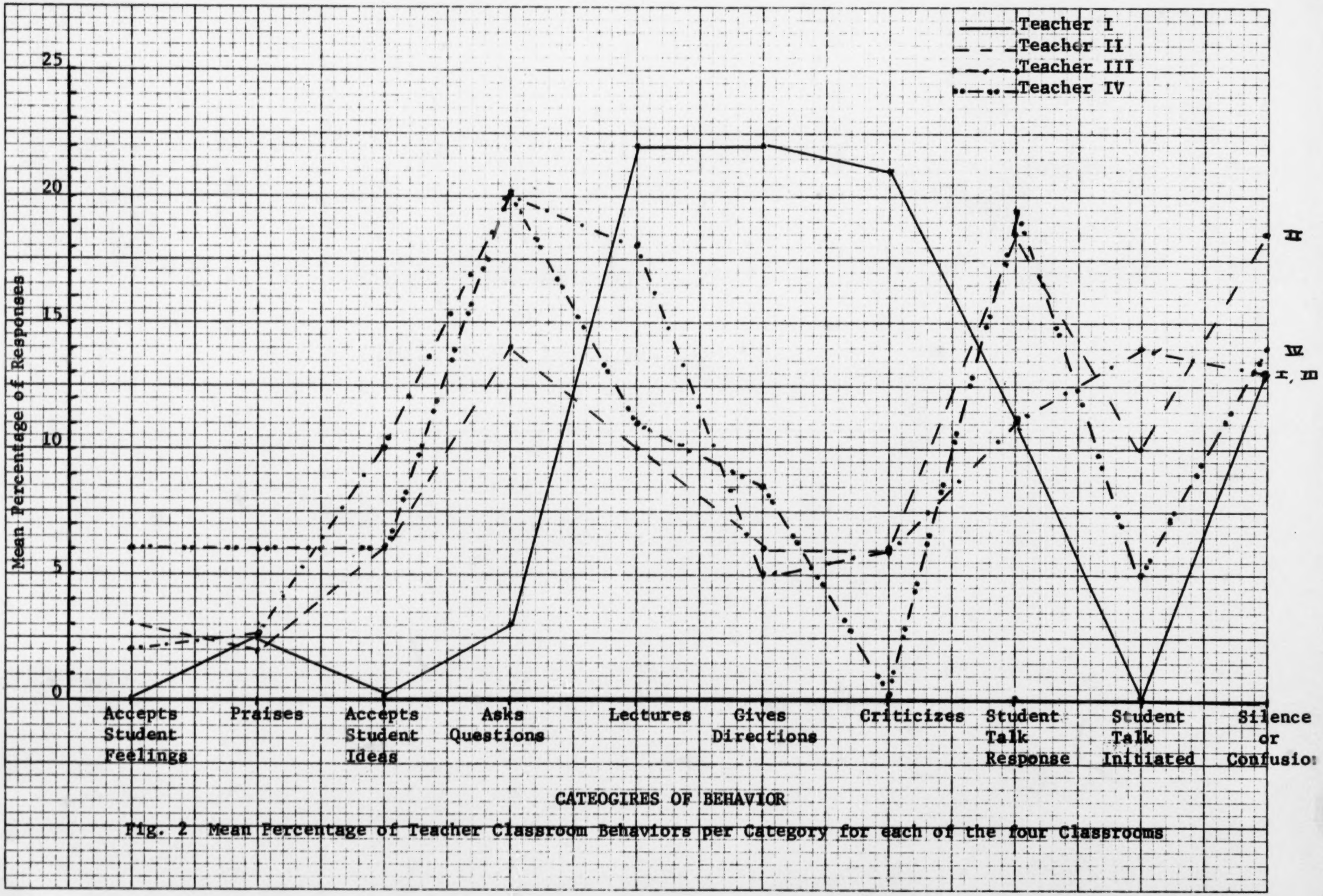

Fig. 2 Mean Percentage of Teacher Classroom Behaviors per Category for each of the four Classrooms